Social Emotion Mining: An Insight

Shivangi Chawla, Monica Mehrotra

Department of Computer Science, Jamia Millia Islamia University, New Delhi, India

Abstract— Emotions are an indispensable component of variety of texts present on online social media services. A lot of research has been done to detect and analyse the emotions present in text but most of them are done from the author's perspective. This paper focuses on providing an in-depth survey of different work done in Social Emotion Mining (SEM) from reader's perspective. It is a first attempt towards categorization of existing literature into emotion mining levels. It also highlights different models and techniques utilized by various authors in this area. Major limitations and challenges in this area of Emotion Detection and Analysis are also presented.
Keywords— SEM, Single-label Classification, Multi-label Classification, readers 'emotion.

I. INTRODUCTION

Social Media Services have become a part and parcel of our lives, through which people are expressing their opinions. The opinion of the public is a reflection of the individual sentiments and emotions. Emotions are contributing in bringing people together of similar likes and dislikes and also have great impact on others. Sentiment analysis and emotion mining of such opinionated text has gathered much attention of the researchers, yet most of the work is done with the intent of finding emotions of the authors. However, writing is an act of communication directed at a target. On the Internet, Writing –even on personal blogs –is aimed at an audience, whether a single person (e.g., instant messaging), a group of individuals (e.g., Facebook), or anyone (e.g., public discussion forums) [1]. In many cases, the reader's emotions triggered by the documents do not always agree with those of writers [2]. Hence, mining emotions from reader's perspective is more meaningful and is termed as Social Emotion Mining. It attempts to bridge the gap between social media materials and readers' emotions [6].

Many online news services now provide a service that lets users convey their emotions after browsing a news article, and then records the emotional responses shared by readers from a set of predefined emotion labels incrementally. The aggregation of such emotional responses is known as social emotions [9]. Mining readers' emotions from textual content is considered to be a predictive mining task through which evoked emotions of readers are foreseen and analyzed. Grasping such emotions can help understand the perspectives and preferences of individual users, and thus may facilitate online publishers to provide more personalized services or statistically study readers' attitudes towards social events [4].

Detecting readers' emotions will assist the authors in foreseeing how their work will influence the emotions of readers. Some potential social media applications include emotion-based document retrieval and emotion aware news recommendation [15]. It can also prove valuable to gain insight into public opinion such as helping enterprises gauge the attitudes of the public to certain exposed incidents so that they can take correct remedial actions [5].Unfortunately, the readers' emotions has not attracted much attention. Detecting emotions from reader's perspective proves to be a challenging task as the readers' emotions are not present in the text rather they are triggered by reading the text. Research on this topic is relatively sparse as compared to those from writers' point of view [16].

II. RELATED WORK

Text can be mined for emotions from two different perspectives: From the perspective of an author/writer and reader. Former studies the emotional intent of the author/writer of the text and the latter tries to identify emotions that are evoked in a reader as a response to the text readership. The writer may directly express her feelings through some emotional words or emoticons. However, the reader's emotion can be invoked by not only the received content but also her personal experiences or knowledge [7]. The author of certain text may intentionally/unintentionally insert certain emotions and opinions towards the facts by omitting or stressing upon some aspect of text, however the interpretation of text largely depends on the personal background knowledge and perception of reader. Due to this emotion bias, only emotions of readers can accurately represent the categories of text like online news. Conventional Emotion mining research focused primarily on detecting the emotions of authors who created the text, while few focus on predicting the impact of reading a particular textual content on the reader. Recognizing emotions from text is considered as a classification task where a particular text is classified into Single/Multiple emotion categories.

First piece of research in this direction was "affective text" in SemEval-2007 Tasks which provided annotation to news headlines based on the emotions they evoked in the minds of reader where a corpus of news headlines extracted from Google-News and CNN was provided to explore the connection between evoked emotions of the readers and the news headlines. The assumption behind this work was that every word, even those apparently neutral, can evoke pleasant or painful experiences due to their semantic relation with emotional concepts or categories [10]. A similar piece of research work had studied readers' emotional states evoked by news sentences [7, 8].The existing approaches to social emotion mining can be classified into two categories: Single-label classification and Multi-label classification. Past works also suggest that research in this field can be done at two levels: word level and topic-level [3]. Both approaches have their own findings and a variety of models have been proposed based on each of them, however topic-level approaches prove to be more efficient as words when associated with certain topics, more accurately convey emotions that a reader might get induced with.

2.1   Social Emotion Mining Techniques

The problem of Social Emotion Detection is mostly studied as a classification problem by past researchers. A piece of text can be classified into one or more emotion categories based on the type of technique used. However, preliminary research in the field of Social Emotion mining of Affective text has taken it as a single label classification problem. However, this is an inconformity with the reality as readers might get induced with a number of emotions at the same time [2]. For example, a sentence "IPhone is a very good phone but only for those who can afford it", might trigger emotions of disappointment, sadness at the same time.

2.1.1   Single-Label Classification

Emotions can be easily classified into one or more categories such that one emotion can be easily distinguished from another category .Reader's emotion detection from text is a data mining process in which a piece of text is classified into an emotion category based on the emotion the reader might get induced with. Major work in this area has focussed on the single-label classification techniques where the affective text is classified into one emotion category including SemEval-2007 tasks. Lin et al. [9] studied Chinese texts for reader emotion prediction and collected data from Yahoo! and Kimo News to classify text into eight emotion categories. They carried out two aspects of experiments one was single label classification and other was ranking of reader emotions [3].

2.1.2   Multi-Label Classification

Classifying text into single emotion category that the reader is predicted to get evoked with is inconsistent with the reality as text may trigger multiple emotions into a single reader. Moreover, treating reader emotion prediction as a single-label classification problem also makes the improvement of classification performance become hard due to the limitations of emotion model [2]. Bhowmick et al. [10] classified the news sentences into multiple emotion categories. Being similar to the task of SemEval-2007, the influence of indirect emotional words was also taken into consideration. Lu ye et al. [2] investigated various multi-label classification algorithms and feature selection methods on a large-scale corpora and investigated that RAkEL performs best with feature selection based on chi-square statistics and document frequency. Y Rao et al. [6] also proposed Multi-label Supervised Topic model which outperforms Emotion-Topic Model. Y. Zhang et al. proposed a supervised topic model for Multi-label emotion tagging for online news [4]. Supervised learning through multi-label topic models treat training documents uniformly and the documents that evoke prominent emotions in users are usually mixed with noisy documents that do not convey much affective meaning. To alleviate this problem, weighing of training documents is done. Li et al. [22] presented a reader perspective weighted model(RPWM) for social emotion classification over varied scale training documents with emotional entropy concept for the estimation of different training documents.

2.2   Social Emotion Mining Levels

Mining textual content for emotions is a challenging text mining task which can be done at various levels of semantics such as at individual word level which give rise to sentences and further into documents and at topic level. Apart from this, each text is associated to a particular context in which its words attain intended meaning, while this meaning changes in another context. So, Social emotion Mining can also be performed at context level.

2.2.1   Word-Level

Words are smallest lexical units of a sentence that can convey any emotion. Each and every word conveys single/multiple emotions in the minds of readers. Words lay the foundation of a sentence by combining themselves and hence describe how users understand and express emotions. Taking this as baseline, most preliminary works detect readers' emotions by exploiting the semantics of individual words, thereby analysing the emotions hidden behind the words. The "affective text" in SemEval-2007 Tasks also focused on word-level approach [10]. In SWAT system, a word emotion mapping dictionary has been first constructed, in which, each word is scored according to multiple emotion labels,

such as joy, fear, anger, surprise etc. Then the dictionary is used to detect social emotions of unlabelled news headlines. An Emotion-Term Model was also developed as a variant of Naive Bayes, to model the word-emotion associations by exploiting all words in the news content.

However, the underlying assumption that "word" is the only essential feature in social emotion mining suffers from many problems. One is sentiment ambiguity in different topics or contexts, as the same word may reflect different emotions. Another problem is due to the background noisy words [6]. Treating each word independently for constructing word-level emotion lexicon may mix background noisy words with relevant words, hence, actual emotions(s) are not conveyed.

2.2.2   Topic-Level

Words are generally associated with different topics i.e. same word may convey different meanings when related to different topics. Topic here refers to any object, entity or real world event that indicates the subject or context of the sentiment. To solve the problems faced in word-level approach, topic models have been developed like Emotion Topic model (ETM) which introduced an intermediate layer into LDA, in which a topic acts as an important component of an emotion [9]. These were single-label topic models. However, some multi-label topic models have also been proposed.

A similar piece of research also presents two sentiment topic models from the readers' perspective, namely the Multi-label Supervised Topic Model (MSTM) and the Sentiment Latent Topic Model (SLTM). To differentiate between affective and background topics, yet another model known as Affective Topic Model (ATM) has been proposed. It also allows differentiating between the different senses of the same word [7].All the above models are joint emotion-topic models as they include additional layer into LDA.

Social Emotion mining by classifications without considering topic in mind, are however useful and effective but they do not find which aspects the general public like or dislike. Without knowing the targets of each emotional state, the classified emotions are of limited use [15]. So, Topic modelling is favoured as it focuses on finding meaningful latent topics from documents.

Apart from these techniques and models, some systems for Social Emotion Analysis and detection have also been proposed, like Peng Nie et al. [19] presented a Social Emotion Analysis System (SEAS) by utilizing Naïve Bayes, logistic regression and ETM to classify the news articles to an emotion category. A yet another System for detecting social emotions was proposed by Lei et al. [15].

2.2.3   Context-Level

Despite the growing amount of work in this research area, adaptive social emotion classification remains a problem. Most works on social emotion detection, including ETM, ATM, MSTM and SLTM, have focused mainly on associating emotions with topics specific to one context.

Context level emotion mining trains a model on one source context and uses it to build model for another target context. Context-sensitivity of topics makes it a challenging task. Many algorithms have been proposed to tackle this problem of context-sensitivity from the perspective of writer. A recent piece of work has addressed the problem of Adaptive social emotion classification from the perspective of reader and built a contextual sentiment topic model (CSTM) to classify reader emotions across different contexts.

Literature reviewed with notion of existing models for emotion mining has been summarized in table 1.

Table 1: Emotion Mining Models

Author	Data Set	Emotion Model	References
Yangui Rao et al.	Online News	Affective Topic Model(ATM)	[7]
Yangui Rao et al.	Online News	MSTM,SLTM	[6]
Rao et al.	Online News	CSTM	[8]
Zhang et al.	Online News	Multi-label supervised emotion topic model(ML-sETM)	[3]
Bao et al.	Online News	Emotion Topic Model(ETM)	[11]
Bao et al.	Online News	Emotion-Term Model(ET)	[11]
X Li et al.	Online news	Weighted Multi-label Classification Model(WMCM)	[4]
Katz et al.	Online News	SWAT	[10], [11]
Li et al.	-	Reader Perspective Weighted Model(RPWM)	[22]

III.   CONCLUSION

On the whole, it is clear that Social Emotion Mining is helpful for understanding the preferences and perspectives of online readers, and therefore can facilitate

the provision of more relevant and personalized services to users by news websites, including emotion-based document retrieval and classification of online news articles by emotion. Several techniques have been proposed for this task based on classification and topic modelling, based on which a number of models have been proposed. There is also a need to investigate how emotional contagion is achieved and how different emotions affect others' emotions by propagating themselves. It will help organisations to shape up their text according to certain emotions which have desired propagation effect.

IV. CHALLENGES AND FUTURE DIRECTIONS

Social emotion mining proves to be a challenging task which faces a number of challenges at present. Firstly due to rapid development of web there is a scarcity of large-scale corpora, which fades recently. Secondly, emotion mining research as multi-label classification still requires extensive improvements. Thirdly, there is a lack of generalization in this area of research as major piece of research has taken online news texts into consideration for mining reader emotions, others have considered social networking sites, but the results of any single model are not proved yet, for all types of texts. So, in future these aspects may be taken into research. Apart from it, there is a demand of studying how emotions propagate online and impact others emotional states developing into an emotional contagion.

REFERENCES

[1] Paltoglou, G., Theunis, M., Kappas, A., &Thelwall, M. (2013). Predicting emotional responses to long informal text. *Affective Computing, IEEE Transactions on*, *4*(1), 106-115.

[2] Ye, L., Xu, R. F., &Xu, J. (2012, July). Emotion prediction of news articles from reader's perspective based on multi-label classification. In *Machine Learning and Cybernetics (ICMLC), 2012 International Conference on* (Vol. 5, pp. 2019-2024). IEEE.

[3] Zhang, Y., Su, L., Yang, Z., Zhao, X., & Yuan, X. (2015). Multi-Label Emotion Tagging for Online News by Supervised Topic Model. In *Web Technologies and Applications* (pp. 67-79). Springer International Publishing.

[4] Li, X., Xie, H., Rao, Y., Chen, Y., Liu, X., Huang, H., & Wang, F. L. (2016, January). Weighted multi-label classification model for sentiment analysis of online news. In *2016 International Conference on Big Data and Smart Computing (BigComp)* (pp. 215-222). IEEE.

[5] Quan, X., Wang, Q., Zhang, Y., Si, L., &Wenyin, L. (2015). Latent Discriminative Models for Social Emotion Detection with Emotional Dependency. *ACM Transactions on Information Systems (TOIS)*, *34*(1), 2.

[6] Rao, Y., Li, Q., Mao, X., &Wenyin, L. (2014). Sentiment topic models for social emotion mining. *Information Sciences*, *266*, 90-100.

[7] Rao, Y., Li, Q., Wenyin, L., Wu, Q., &Quan, X. (2014). Affective topic model for social emotion detection. *Neural Networks*, *58*, 29-37.

[8] Rao, Y. (2016). Contextual Sentiment Topic Model for Adaptive Social Emotion Classification. *Intelligent Systems, IEEE*, *31*(1), 41-47.

[9] Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., & Yu, Y. (2012). Mining social emotions from affective text. *Knowledge and Data Engineering, IEEE Transactions on*, *24*(9), 1658-1670.

[10] Strapparava, C., &Mihalcea, R. (2007, June). Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 70-74). Association for Computational Linguistics.

[11] Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., & Yu, Y. (2009, December). Joint emotion-topic modeling for social affective text mining. In*Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on* (pp. 699-704). IEEE..

[12] Lei, J., Rao, Y., Li, Q., Quan, X., &Wenyin, L. (2014). Towards building a social emotion detection system for online news. *Future Generation Computer Systems*, *37*, 438-448.

[13] Ferrara, E., & Yang, Z. (2015). Measuring Emotional Contagion in Social Media. *PloS one*, *10*(11), e0142390.

[14] Nie, P., Zhao, X., Yu, L., Wang, C., & Zhang, Y. (2015, September). Social Emotion Analysis System for Online News. In *2015 12th Web Information System and Application Conference (WISA)* (pp. 43-48). IEEE.

[15] Yang, C., Lin, K. H. Y., & Chen, H. H. (2009, September). Writer Meets Reader: Emotion Analysis of Social Media from both the Writer's and Reader's Perspectives. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 01* (pp. 287-290). IEEE Computer Society

[16] Tang, Y. J., & Chen, H. H. (2011, November). Emotion modeling from writer/reader perspectives using a microblog dataset. In *Proceedings of IJCNLP Workshop on sentiment analysis where AI meets psychology* (pp. 11-19).

[17] Tang, Y. J., & Chen, H. H. (2012). Mining Sentiment Words from Microblogs for Predicting Writer-Reader Emotion Transition. In *LREC* (pp. 1226-1229).

[18] Bhowmick, P. K. (2009). Reader perspective emotion analysis in text through ensemble based multi-label classification framework. *Computer and Information Science*, *2*(4), 64.

[19] Li, W., &Xu, H. (2014). Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications*, *41*(4), 1742-1749.

[20] Bhowmick, P. K., Basu, A., Mitra, P., & Prasad, A. (2009, December). Multi-label text classification approach for sentence level news emotion analysis. In *International Conference on Pattern Recognition and Machine Intelligence*(pp. 261-266). Springer Berlin Heidelberg.

[21] Li, X., Rao, Y., Chen, Y., Liu, X., & Huang, H. (2016, March). Social emotion classification via reader perspective weighted model. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence, poster paper.*

[22] Rao, Y., Xie, H., Li, J., Jin, F., Wang, F. L., & Li, Q. (2016). Social emotion classification of short text via topic-level maximum entropy model. *Information & Management.*

[23] Yu, L., Zhao, X., Wang, C., Zhang, H., & Zhang, Y. (2016, June). Cross-Collection Emotion Tagging for Online News. In *International Conference on Web-Age Information Management* (pp. 225-237). Springer International Publishing.

[24] Lin, K. H. Y., Yang, C., & Chen, H. H. (2007, July). What emotions do news articles trigger in their readers?. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 733-734). ACM.